

Working Paper 97-26
Statistics and Econometrics Series 13
February 1997

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (341) 624-9849

NONLINEAR ERROR CORRECTION MODELS.

Alvaro Escribano and Santiago Mira*

Abstract

The relationship between cointegration and error correction models (EC) is well characterized in a linear context, see Engle and Granger (1987) and Johansen (1991), but the extension to the nonlinear context is still a challenge. Few extensions of the linear framework were done in the context of nonlinear error correction (NEC), see Escribano (1986 and 1987), or asymmetric and time varying error correction models, see Granger and Lee (1989) and Burgess (1992). In this paper we propose a theoretical framework based on the concept of near epoch dependence (NED) that allow us to formally address those issues. In particular, we partially extend Granger Representation Theorem to the nonlinear case and we study the estimation and inference properties of least squares when the cointegrating relation is linear but the dynamic model is a NEC. The two-step estimation approach of Engle and Granger (1987) is extended when the cointegrating errors are NED and the dynamic model is a NEC. Some potentially useful NEC models are proposed and Monte Carlo simulations are provided.

Keywords:

Cointegration, nonlinear error correction, near epoch dependence, two-step least squares estimator.

*Escribano, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid. C/ Madrid, 128 28903 Madrid. Spain. Ph: 34-1-624.98.54, Fax: 34-1-624.98.49, e-mail: alvaroe@est-econ.uc3m.es; Mira, Departamento de Fundamentos del Análisis Económico, Universidad Europea de Madrid. The authors acknowledge financial support from spanish DGICYT PB95-0298.

1 Introduction

Granger (1981) introduced the concept of cointegration but it was not until Engle and Granger (1987) and Johansen(1988, 1991) that this concept got an immense popularity among econometricians and applied economists. The great impact those papers had in the profession was due to the fact that they showed how we should statistically work with economic variables that are non stationary, in order to avoid the problem of spurious regressions, see Granger and Newbold (1974) and Phillips (1986). Furthermore, most of the estimation and inference procedures changed dramatically from the classical statistical frameworks when dealing with variables that have unit roots and are cointegrated. By now it is clear how to deal with integrated and cointegrated data in a linear context, see Watson (1996), but almost no research has been dedicated to the simultaneous consideration of nonstationarity and nonlinearity. Even though many economist agree that those are dominant and likely properties of many economic data. How can it be possible that only few research have been dedicated to this topic? The answer is clear, it is difficult to work with nonlinear time series models within a stationary and ergodic framework and therefore even more difficult within a nonstationary context.

An introduction to the state of the art in econometrics relating nonlinearity and nonstationarity within a time series context can be found in Granger and Teräsvirta (1993) and Granger (1995). Those authors discussed the concepts of long-range dependence in mean and extended memory which generalize the linear concept of integration, $I(1)$, to a nonlinear framework. The main disadvantage of those definitions is that they have no Laws of Large Numbers (LLN), nor Functional Central Limit Theorems (FCLT) associated to them, and therefore its hard to obtain estimation and inference results. On the other hand, there are interesting empirical macroeconomic applications where nonlinearity has been found in a nonstationary context and therefore, there is a need to econometrically justify those results. This paper starts filling this mayor gap.

As an empirical application of nonlinear error correction (NEC) models we have the case of the UK money demand from 1878 to 1970. With this data set, Escribano (1986) showed that a nonlinear error correction was the best specification of the alternatives linear models that were proposed by Friedman and Schwarz (1982), Hendry and Ericsson (1985) and Longbottom and Holly (1985). In fact in their latest specification of that money demand for the UK, Hendry and Ericsson (1991) used the nonlinear error correction specification. The variables are: m =log money stock (millions), i =log real net national product, p =deflator of i , rs =log of short term interest rate, rl =log of long-term interest rate, and RS =short term interest rate. L is the lag operator such that $L^k x_t = x_{t-k}$. Let \hat{u}_t be the residuals from the cointegrating relationship estimated by OLS, then the two-step approach of Engle and Granger(1987) is given by:

$$\hat{u}_t = (m - p - y)_t + 0.309 + 7RS_t$$

$$\begin{aligned}
(1-L)(m-p)_t = & 0.45(1-L)(m-p)_{t-1} - (1-L)^2(m-p)_{t-2} - 0.60(1-L)p_t + \\
& 0.39(1-L)p_{t-1} - 0.021(1-L)rs_t - 0.062(1-L^2)rl_t - \\
& 2.55(\hat{u}_{t-1} - 0.2)\hat{u}_{t-1}^2 + 0.005 + 3.7(D1 + D3) + \varepsilon_t
\end{aligned}$$

In the second step, the term \hat{u}_{t-1} enters in the nonlinear error correction term, and we can estimate the rest of the parameters of the dynamic formulation by OLS. This error correction model is nonlinear since the nonlinear adjustment is a cubic polynomial. The sign of the estimated coefficient (-2.55) guarantees that the adjustment is in fact error correcting. In this paper we give sufficient conditions for those OLS parameters to be consistently estimated and we show that the short run parameters have a limiting distribution that is normal. Other empirical examples of nonlinear error correction models are Granger and Lee (1989), Balke and Fomby (1992), Burgess (1992), Kunst (1992) and Granger and Swanson (1995). The main goal of this paper is to give sufficient conditions for the parameters of those models to be consistently estimated by two-step least squares and to show that the short run parameters of the dynamic model have a limiting normal distribution.

The structure of this paper is the following. In Section 2 we propose an alternative concept of integration, $I(0)$ and $I(1)$, which can also be extended to nonlinear cointegration. Based on those definitions we study if certain results of Granger Representation Theorem are maintained in this nonlinear framework. We propose a representation theorem which relates the concept of linear cointegration with the nonlinear error correction, introduced by Escribano (1986, 1987). Section 3, deals with the least squares estimation of linear cointegrating relationships in a general context with cointegrating errors that are near epoch dependent. We derived the asymptotic distribution of the two-step least squares estimator of the short run parameters of the NECs. Section 4, covers some Monte Carlo simulations to study the small sample properties of least squares in several parametric NEC models. Finally in section 5, we present the main conclusions.

2 Nonlinear Error Correction Models

2.1 Definitions

A general concept of $I(0)$ for a sequence $\{v_t\}$ is given by the "high level" condition that v_t verifies a Functional Central Limit Theorem (FCLT), i.e. that $T^{-1/2} \sum_{t=1}^{[Tr]} v_t \xrightarrow{d} B(r)$ where $B(r)$ is a Brownian Motion. In a (non)-linear dynamic model this FCLT holds for functions of the exogenous variables and underlying disturbances that have a sufficiently fading memory. Different approaches to modelize these dynamics have been developed: Bierens (1981) employ the concept of "stochastically stable" w.r.t. an α -mixing sequence:

Gallant and White (1988) or Wooldridge and White (1988) employ the concept of "near epoch dependence" (NED) w.r.t. an α -mixing sequence. Both concepts need to assume that the exogenous variables and disturbances are α -mixing (therefore sequences asymptotically independent) to provide useful results. The definition of $I(0)$ that we are going to use is based in the concept of NED.

Definition 2.1a (NED) Let $\{z_t : \Omega \rightarrow \mathbb{R}\}$ be a sequence $(\mathcal{F}, \mathcal{B})$ -medible with $E(z_t^2) < \infty$ for all t . Then it will be said that $\{z_t\}$ is near epoch dependent (NED) on the underlying sequence v_t iff $\{\phi_m\}$ is of size $-a$, for ϕ_m given by

$$\phi_m \equiv \sup_t \|z_t - E_{t-m}^{t+m}(z_t)\|_{L_2}$$

and where $E_{t-m}^{t+m}(z_t) = E(z_t | v_{t-m}, \dots, v_{t+m})$ and $\|\cdot\|_{L_2}$ is the norm L_2 of a random variable, defined as $E^{1/2}|\cdot|^2$.

We assume that the future values of v_t do not improve the conditional expectation of z_t , in the sense of Sims (1972), such that the forward values v_{t+r} ($r = 1, \dots, m$) are useless, but harmless. From Definition 2.1a we can say that ϕ_m is the worst mean square forecast error when z_t is predicted by $E_{t-m}^{t+m}(z_t)$. When ϕ_m goes to zero at an appropriate rate, then z_t depends essentially on the recent epoch of v_t . If z_t depends on a finite number of lags of v_t then it is NED of any size. The property of NED is maintained under sums and products (see Gallant and White (1988)) and under some conditions verifies a Law of Large Numbers (LLN) and a FCLT. A simplified version for arrays of that FCLT is as follows.

Theorem 2.1b (FCLT for NED, see Wooldridge and White (1988)) Consider the following assumptions: (i) $\{z_{nt} : n, t = 1, 2, \dots\}$ is a double array of real-valued random variables; (ii) $\sigma_n^2 = \text{Var}(\sum_{t=1}^n z_{nt})$ verifies that $\{\sigma_n^{-2}\}$ is $O(n^{-1})$; (iii) for some $r > 2$ and all n, t : $\|z_{nt}\|_{L_2} < \Delta < \infty$ and $E(z_{nt}) = 0$; (iv) $\{z_{nt}\}$ is NED w.r.t. $\{y_{nt}\}$ of size $-1/2$; (v) $\{y_{nt}\}$ is α -mixing of size $-r/(r-2)$; (vi) for each $r \in [0, 1]$, $E(W_n(r)^2) \rightarrow r$ as $n \rightarrow \infty$, where the sequence $\{W_n(\cdot)\}$ is defined as

$$W_n(r) = \sum_{t=1}^{\lfloor nr \rfloor} \sigma_n^{-1} z_{nt}.$$

Then, $\{W_n\}$ verifies $W_n \xrightarrow{d} W$, where W is the Standard Brownian Motion.

The above considerations motivates the following definition.

Definition 2.2 A sequence $\{\varepsilon_t\}$ is $I(0)$ if it is NED on an underlying α -mixing sequence $\{v_t\}$ but the sequence $\{x_t\}$ given by $x_t = \sum_{s=1}^t \varepsilon_s$ is not NED. We will say that x_t is $I(1)$.

Notice that if x_t is $I(1)$ then Δx_t is $I(0)$. This definition of $I(0)$ plus the conditions of Theorem 2.1b ensure a FCLT for the $I(0)$ series.

Definition 2.3 Two sequences $\{y_t\}$ and $\{x_t\}$ which are $I(1)$ are cointegrated with cointegration function $g(\cdot, \cdot; \gamma)$, if $g(y_t, x_t; \gamma^*)$ is NED on some α -mixing sequence but the sequence $g(y_t, x_t; \gamma)$, is not NED for $\gamma \neq \gamma^*$.

This definition allows us to extend the notion of cointegration to a nonlinear context without having to deal with the difficulties faced by Escribano (1987) or Granger and Hallman (1991) when characterizing the time series properties of nonlinear transformations of series that are $I(0)$ or $I(1)$. It also allows us to deal formally NEC models.

Definition 2.4 A non linear error correction (NEC) mechanism for the $(n \times 1)$ X_t vector is an autoregressive linear model (VAR) for the differences ΔX_t plus nonlinear terms for the lag of the levels, say X_{t-1} .

If we take the case $n = 2$ and $X_t = [x_t, y_t]'$, the NEC with only one lag is

$$\Delta X_t = \Psi_1 \Delta X_{t-1} + H(X_{t-1}; \Gamma) + \eta_t,$$

whose first equation can be written in the form

$$\begin{aligned} \Delta y_t &= \psi_{11} \Delta y_{t-1} + \psi_{12} \Delta x_{t-1} + h(y_{t-1}, x_{t-1}; \gamma) + \varepsilon_{1t} \\ \Delta y_t &= \psi_{11} \Delta y_{t-1} + \psi_{12} \Delta x_{t-1} + h(g(y_{t-1}, x_{t-1}; \gamma_1); \gamma_2) + \varepsilon_{1t} \end{aligned} \quad [2.1]$$

where Δy_t and Δx_t are NED, and the parameter γ may be split into $\gamma = [\gamma'_1, \gamma'_2]'$. The subvector γ_1 is the cointegration vector and the subvector γ_2 is the vector of parameters of the nonlinear error correction mechanism. Notice the distinction made in [2.1] between the cointegration function $g(y_t, x_t; \gamma_1)$ and the error correction function $h(\cdot; \gamma_2)$. The function $g(\cdot, \cdot; \gamma_1) = 0$ gives the long run equilibrium relationship and the deviations from this equilibrium $g(y_{t-1}, x_{t-1}; \gamma_1)$ are the errors corrected by $h(\cdot; \gamma_2)$.

In order to study NEC models with linear cointegration [2.1] becomes

$$\begin{aligned} \Delta y_t &= \psi_{11} \Delta y_{t-1} + \psi_{12} \Delta x_{t-1} + h(z_{t-1}; \gamma_2) + \varepsilon_{1t} \\ z_{t-1} &= y_{t-1} + \gamma_1 x_{t-1} \equiv g(y_{t-1}, x_{t-1}; \gamma_1) \end{aligned} \quad [2.2]$$

Consider the following example. Let $X_t = [y_t, x_t, r_t]'$, and $H(X_{t-1}) = J(KX_{t-1})$ for some matrix K ($n \times n$) and some function $J : \mathbb{R}^n \rightarrow \mathbb{R}^n$, and suppose that K is given by linear combinations of two cointegrating relations, i.e.

$$K = \begin{pmatrix} K'_1 \\ K'_2 \\ K'_3 \end{pmatrix} = \gamma \alpha' = \begin{pmatrix} \gamma'_1 \\ \gamma'_2 \\ \gamma'_3 \end{pmatrix} \begin{pmatrix} \alpha'_1 \\ \alpha'_2 \end{pmatrix} = \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \\ \gamma_{31} & \gamma_{32} \end{pmatrix} \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \end{pmatrix}$$

It is clear that the function of $(K'_1 X_{t-1}, K'_2 X_{t-1}, K'_3 X_{t-1})$ can be written as a function of only $[\alpha'_1 X_{t-1}, \alpha'_2 X_{t-1}]' = [z_{1,t-1}, z_{2,t-1}]'$. Therefore, the error correction model with only one lag is given by

$$\begin{aligned}\Delta y_t &= \beta_1 \Delta y_{t-1} + \beta_2 \Delta x_{t-1} + \beta_3 \Delta r_{t-1} + J_1(z_{1,t-1}, z_{2,t-1}) + \varepsilon_{1t} \\ \Delta x_t &= \delta_1 \Delta y_{t-1} + \delta_2 \Delta x_{t-1} + \delta_3 \Delta r_{t-1} + J_2(z_{1,t-1}, z_{2,t-1}) + \varepsilon_{2t} \\ \Delta r_t &= \rho_1 \Delta y_{t-1} + \rho_2 \Delta x_{t-1} + \rho_3 \Delta r_{t-1} + J_3(z_{1,t-1}, z_{2,t-1}) + \varepsilon_{3t}\end{aligned}$$

where the error correction is a function of the base of cointegrating relations.

In the next section we provide a partial generalization of the Granger Representation Theorem given in Engle and Granger (1987).

2.2 A Representation Theorem

Before characterizing the representation theorem it is convenient to introduce some results that will be instrumental in the proof. Consider the following model

$$Z_t = \Phi_1 W_{t-1} + F(Z_{t-1}; \gamma) + U_t \quad [2.3]$$

where Z_t and U_t are $(r \times 1)$, W_t is $(n \times 1)$, Φ_1 is $(r \times n)$, and $F(\cdot; \gamma) : \mathbb{R}^r \rightarrow \mathbb{R}^r$ as a function of Z . Assumption 2.5 and Theorem 2.6 that follow will be useful later on.

Assumption 2.5

- (a) The sequence $\{U_t\}$ is α -mixing of size $-v/(v-2)$ for $v > 2$, and the sequence $\{W_t\}$ given in [2.3] is NED on an underlying α -mixing sequence $\{A_t\}$, of size $-v/(v-2)$ for $v > 2$, in the sense that for v_m given as

$$v_m \equiv \sup_t E \|W_t - E(W_t | A_t, \dots, A_{t-m})\|_S^2$$

it holds that $v_m \rightarrow 0$ as $m \rightarrow \infty$, where the norm $\|\cdot\|_S$ is introduced in Mira and Escribano (1995). See Appendix A.

- (b) For the norm $\|\cdot\|_S$ we have

$$\|\nabla_Z F(Z; \gamma)\|_S \equiv \delta_Z < 1.$$

- (c) The following moment conditions hold for $i = 2$

- (i) $E\|W_t\|_S^i \leq \Delta_W^{(i)}$,
- (ii) $E\|U_t\|_S^i \leq \Delta_U^{(i)}$,
- (iii) $E\|U_t\|_S^i \|W_t\|_S^i \leq \Delta_{WU}^{(i)}$.

(d) $F(\cdot; \gamma)$ is continuously differentiable in each argument.

If in Assumption 2.5 (a) U_t is NED instead of α -mixing, then the proof would change only slightly but the result is valid. Assumption 2.5 (b) says that the spectral radius of the matrix of first partial derivatives is smaller than 1. We will see later on that this assumption plays an important role.

Theorem 2.6 Under Assumption 2.5 the sequence $\{Z_t\}$ given in [2.3] is NED on the underlying α -mixing sequence $\{(U_t, A_t)\}$ of any size. \square

Proof: See Appendix A.

The core of the proof is that if Z_t is NED on W_t and W_t is NED on A_t then Z_t is NED on A_t . Now we have the tools to give a representation theorem for a nonlinear error correction with linear cointegration, in the sense that we give sufficient conditions to ensure a balanced specification of NEC models.

Theorem 2.7 (Representation Theorem) Consider the following nonlinear time series model for the sequence of $(n \times 1)$ vectors $\{X_t\}$, given by

$$X_t = F(X_{t-1}, X_{t-2}) + \varepsilon_t \quad [2.4]$$

where for simplicity only two lags are included. Suppose the following conditions:

- (1) ε_t is α -mixing and ΔX_t is $I(0)$;
- (2) the function $F(X_{t-1}, X_{t-2})$ is nonlinear only in the first lag, i.e.

$$F(X_{t-1}, X_{t-2}) = G(X_{t-1}) + \Phi_2 X_{t-2};$$

and

- (3) $H(X_{t-1}) = J(\alpha' X_{t-1})$ where $\alpha' X_{t-1} \equiv Z_{t-1}$.

Then

- (i) under Assumption (2) we have the following representation

$$\Delta X_t = \Psi_1 \Delta X_{t-1} + H(X_{t-1}) + \varepsilon_t \quad [2.5]$$

where $\Psi_1 = -\Phi_2$ and $H(X_t) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is given by $H(X_{t-1}) = -(I - \Phi_2)X_{t-1} + G(X_{t-1})$;

(ii) under Assumption (3), if we multiply [2.5] by α' we obtain the following representation for Z_t ,

$$Z_t = \Phi_1 W_{t-1} + F(Z_{t-1}) + U_t \quad [2.6]$$

where $Z_t = \alpha' X_t$, $W_t = \Delta X_t$, $\Phi_1 = \alpha' \Psi_1$, and $F(Z_{t-1}) = \alpha' J(\alpha' X_{t-1}) + \alpha' X_{t-1}$;

(iii) under Assumptions (1)-(3) plus Assumption 2.5 for model [2.6] we have that Z_t is NED.

□

Proof: See Appendix A.

Note that (1) implies on [2.5] that $(1 - \Psi_1 L)$ cannot have a unit root. The result (iii) of the former theorem ensures that under Assumptions (1) to (3) plus 2.6 we have that [2.5] is a correctly specified NEC. If we consider the example given in Section 2.1, the expression [2.6] is given by

$$\begin{aligned} z_{1t} &= \alpha_{11} w_{1,t-1} + \alpha_{12} w_{2,t-1} + \alpha_{13} w_{3,t-1} + z_{1,t-1} \\ &\quad + \alpha_{11} J_1(z_{1,t-1}, z_{2,t-1}) + \alpha_{12} J_2(z_{1,t-1}, z_{2,t-1}) + \alpha_{13} J_3(z_{1,t-1}, z_{2,t-1}) + u_{1t} \\ z_{2t} &= \alpha_{21} w_{1,t-1} + \alpha_{22} w_{2,t-1} + \alpha_{23} w_{3,t-1} + z_{2,t-1} \\ &\quad + \alpha_{21} J_1(z_{1,t-1}, z_{2,t-1}) + \alpha_{22} J_2(z_{1,t-1}, z_{2,t-1}) + \alpha_{23} J_3(z_{1,t-1}, z_{2,t-1}) + u_{2t}. \end{aligned}$$

The condition given by Assumption 2.5 (b) says that $\text{RSpec}(\nabla_Z F(Z)) < 1$ where the function $\text{RSpec}(M)$ is the spectral radius of the matrix M . In this example we have

$$\nabla_Z F(Z) = \begin{pmatrix} 1 + \alpha_{11} \frac{\partial J_1}{\partial z_1} + \alpha_{12} \frac{\partial J_2}{\partial z_1} + \alpha_{13} \frac{\partial J_3}{\partial z_1} & \alpha_{11} \frac{\partial J_1}{\partial z_2} + \alpha_{12} \frac{\partial J_2}{\partial z_2} + \alpha_{13} \frac{\partial J_3}{\partial z_2} \\ \alpha_{21} \frac{\partial J_1}{\partial z_1} + \alpha_{22} \frac{\partial J_2}{\partial z_1} + \alpha_{23} \frac{\partial J_3}{\partial z_1} & 1 + \alpha_{21} \frac{\partial J_1}{\partial z_2} + \alpha_{22} \frac{\partial J_2}{\partial z_2} + \alpha_{23} \frac{\partial J_3}{\partial z_2} \end{pmatrix}.$$

For instance if we have only one equation and only one cointegrating relation then $J_2 = J_3 = 0$ y $z_{2t} = 0$ and the matrix $\nabla_Z F(Z)$ is

$$\nabla_Z F(Z) = \begin{pmatrix} 1 + \alpha_{11} \frac{\partial J_1}{\partial z_1} \end{pmatrix}.$$

Therefore condition $\text{RSpec}(\nabla_Z F(Z)) < 1$ reduces to $|1 + \alpha_{11} \frac{\partial J_1}{\partial z_1}| < 1$. See Mira (1996) for some comments about the case of nonlinear cointegration and nonlinear error correction.

Theorem 2.7 can be as well stated replacing Assumption (1) by Assumption (1') given by

(1') ε_t is $I(0)$:

and in this case we obtain the following theorem. This theorem provides sufficient conditions to jointly ensure that ΔX_t and $\alpha' X_t$ are NED.

Theorem 2.8 Suppose that (1'), (2) and (3) of Theorem 2.7, and Assumption 2.5 apply to model

$$\xi_t = \Pi_1 \xi_{t-1} + \Pi_2 \xi_{t-2} + F(\xi_{t-1}) + \eta_t \quad [2.7]$$

where $\xi'_t = [Z'_t, \Delta X'_{2t}]$, for some partition $X'_t = [X'_{1t}, X'_{2t}]$ that depends on the cointegrating vectors, then ΔX_t and $\alpha' X_t$ are jointly NED.

Proof: see Appendix A.

2.3 Extensions

If the error correction function depends on say two lags X_{t-1} and X_{t-2} , an extension of Theorem 2.7 can be given. Let us write

$$\begin{aligned} X_t &= G(X_{t-1}, X_{t-2}) + \Phi_2 X_{t-2} + \varepsilon_t \\ \Delta X_t &= G(X_{t-1}, X_{t-2}) - X_{t-1} + \Phi_2 X_{t-2} + \varepsilon_t \\ &= (-\Phi_2)(X_{t-1} - X_{t-2}) - (I - \Phi_2)X_{t-1} + G(X_{t-1}, X_{t-2}) + \varepsilon_t \\ &= \Psi_1 \Delta X_{t-1} + H(X_{t-1}, X_{t-2}) + \varepsilon_t \quad [2.8] \end{aligned}$$

where $\Psi_1 = -\Phi_2$ and $H(X_{t-1}, X_{t-2}) = -(I - \Phi_2)X_{t-1} + G(X_{t-1}, X_{t-2})$, and condition (3) of Theorem 2.7 should be changed appropriately. An example of this type of models is the Smooth Transition Regression function (STR) given in Granger y Terasvirta (1993), where the transition depends on some equilibrium errors of the long run relationship specified by the cointegrating relation. For example, if we have $X_t = [y_t, z_t]'$, then the first equation of [2.8] may be written as

$$\begin{aligned} \Delta y_t &= \beta_{11} \Delta y_{t-1} + \beta_{12} \Delta z_{t-1} \\ &\quad + (\delta_{11} \Delta y_{t-1} + \delta_{12} \Delta z_{t-1})(1 + \exp(-\gamma_1(y_{t-1} - \gamma_2 z_{t-1}))) + \varepsilon_{1t} \end{aligned}$$

In this case the dynamics of Δy_t have an autoregressive representation with exogenous variables, whose parameters change depending on the equilibrium errors of the long run relationship.

3 Two Step Estimation

Once the NEC model is correctly specified, it is of interest to give sufficient conditions on the NLS estimator of [2.2] that ensure its consistency and its asymptotic distribution.

Proposition 3.1 (a) Under conditions of a FCLT and if the nonlinear cointegration function is Hadamard differentiable, then the NLS estimator is consistent. (b) Under the conditions given in (a) and some conditions on the derivative of the nonlinear cointegration function, the NLS estimator is superconsistent.

Proof: See Appendix B.

Notice that in the linear case we have the OLS estimator and in that case, the Hadamard differentiability of part (a) and the conditions on the derivative of part (b) hold.

Proposition 3.2 Under conditions of Theorem 2.7 and standard assumptions, [2.2] is a correctly specified model and assuming we know the cointegrating parameters, then the NLS estimators of the rest of the parameters of the NEC model are consistent and asymptotically normal estimators.

Proof: See Appendix B.

For the two step estimation the only thing that remains to prove is that the asymptotic distribution of the estimators given in Proposition 3.2 is the same no matter what value of the cointegrating parameters we use, the estimated values given in Proposition 3.1, say γ_1^T , or the true values, say γ_1^* .

The nonlinear error correction model that we want to estimate is a single equation model with a nonlinear error correction which depends on a single cointegrating relationship, given by

$$\Delta y_t = \beta^* \Delta y_{t-1} + \delta^* \Delta x_{t-1} + f(g(y_{t-1}, x_{t-1}; \gamma_1^*); \gamma_2^*) + v_t$$

which can be written as

$$r_t = \beta^* r_{t-1} + \delta^* w_{t-1} + f(z_{t-1}^*; \gamma_2^*) + v_t \quad [3.1]$$

where $z_{t-1}^* \equiv g(y_{t-1}, x_{t-1}, \gamma_1^*)$ represent the nonlinear equilibrium errors, $\Delta y_t \equiv r_t$, and $\Delta x_t \equiv w_t$. If we stack all the observations in vector form we get

$$\bar{R} - K B^* - F^*(\gamma_2^*) = V \quad [3.2]$$

$$G^*(\theta^*) = V \quad [3.3a]$$

where $\bar{R} = [r_1, \dots, r_T]'$, $R = [r_0, \dots, r_{T-1}]'$, $W = [w_0, \dots, w_{T-1}]'$, $K = [R, W]$, $B^* = [\beta^*, \delta^*]'$, $F^*(\gamma_2^*) = [f(z_0^*; \gamma_2^*), \dots, f(z_{T-1}^*; \gamma_2^*)]'$, and $\theta^* = [\beta^*, \delta^*, \gamma_2^*]'$, and $G^*(\theta^*) = \bar{R} - KB^* - F^*(\gamma_2^*)$.

The two-step estimation procedure proposed by Engle and Granger (1987) consists in estimating the cointegrating parameter in a first step by OLS, say γ_1^T , generate the residuals, and then use those residuals in a second step for estimating the remaining parameters of the nonlinear error correction model [3.1] but substituting z_{t-1}^* by z_{t-1}^T . For instance in a linear case we would substitute $z_t^T = y_t - \gamma_1^T x_t$ for $z_t^* = y_t - \gamma_1^* x_t$. In order to obtain a similar result for the nonlinear case we consider the following assumption.

Define $z_{t-1}^T = g(y_{t-1}, x_{t-1}, \gamma_1^T)$ for γ_1^T the NLS estimation of the cointegrating parameter γ_1^* , and $F^T(\gamma_2^*) = [f(z_0^T; \gamma_2^*), \dots, f(z_{T-1}^T; \gamma_2^*)]'$.

Assumption 3.3 Define the function

$$G^T(\theta^*) = \bar{R} - KB^* - F^T(\gamma_2^*) \quad [3.3b]$$

and assume that the following conditions hold.

$$\lim_{T \rightarrow \infty} (T^{-1} G_\theta^T(\theta^*)' G_\theta^T(\theta^*)) = \lim_{T \rightarrow \infty} (T^{-1} G_\theta^*(\theta^*)' G_\theta^*(\theta^*)) = O_p(1) \quad [3.4]$$

$$\lim_{T \rightarrow \infty} (T^{-1/2} V' G_\theta^T(\theta^*)) = \lim_{T \rightarrow \infty} (T^{-1/2} V' G_\theta^*(\theta^*)) = O_p(1), \quad \text{and} \quad [3.5]$$

$$\lim_{T \rightarrow \infty} (T^{-1/2} (F^*(\gamma_2^*) - F^T(\gamma_2^*))' G_\theta^T(\theta^*)) = o_p(1). \quad [3.6]$$

where $G_\theta^T(\theta^*)$ and $G_\theta^*(\theta^*)$ are the derivatives of the expressions given in [3.3a] and [3.3b].

These assumptions have clear implications in the linear case, see Mira (1996). Consider for instance that $\beta^* = 0$, then $G_\theta^T(\theta^*) = [-W, -F_\gamma^T(\gamma_2^*)]$ and then the first term in [3.4] is the limit of

$$T^{-1} G_\theta^T(\theta^*)' G_\theta^T(\theta^*) = T^{-1} \begin{pmatrix} W'W & W'F_\gamma^T \\ F_\gamma^T F_\gamma^T \end{pmatrix} = \begin{pmatrix} T^{-1} \sum w_t^2 & T^{-1} \sum w_t f_\gamma(z_t^T, \gamma_2^*) \\ T^{-1} \sum (f_\gamma(z_t^T, \gamma_2^*))^2 \end{pmatrix}.$$

If those conditions are difficult to prove analytically, we suggest to do some simulations should be done to ensure that conditions hold. What we found is that the more demanding condition is [3.6] since it affects the nonlinear behaviour of the $f(\cdot)$ function. With the Assumption 3.3 we can prove the following theorem.

Theorem 3.4 Suppose that model [3.1] can be consistently estimated by NLS. Under Assumption 3.3, the estimation of model [3.1] with the cointegration parameter γ_1 estimated by NLS γ_1^T , instead of the true parameter γ_1^* , provides the same asymptotic distribution for

the NLS estimations θ^T of the rest of parameters θ^* , than those θ^T , estimated with the true value γ_1^* .

Proof: See Appendix B.

4 Alternative Nonlinear Error Correction Models

4.1 The Data Generating Process

Consider the NEC with linear cointegration

$$\begin{aligned}\Delta y_t &= \phi_1^* \Delta x_t + f(z_{t-1}^*; \gamma_2^*) + v_t \\ y_t &= \gamma_1^* x_t + z_t^*.\end{aligned}$$

This model is straightforward to generalize to include several variables, lags and cointegrating relations. We have choosen this model in order to simplify the simulations. The data generating process (DGP) is the following. Consider two independent α -mixing sequences $\{a_t\}$ and $\{v_t\}$ with zero mean and define

$$x_t = x_{t-1} + a_t \quad [4.1]$$

$$z_t^* = z_{t-1}^* + \delta_1^* a_t + f(z_{t-1}^*, \gamma_2^*) + v_t \quad [4.2]$$

$$y_t = \gamma_1^* x_t + z_t^* \quad [4.3]$$

where the function $f(z_{t-1}^*, \gamma_2^*)$ is chosen as a parametric nonlinear error correction function. Now if z_t is α -mixing then we obtain that x_t is I(1), y_t is I(1) and they are cointegrated with cointegration function $y_t - \gamma_1^* x_t$, see Theorems 2.7 and 2.8. If we apply the difference operator to [4.3] we obtain

$$\Delta y_t = (\gamma_1^* + \delta_1^*) \Delta x_t + f(z_{t-1}^*, \gamma_2^*) + v_t \quad [4.4]$$

which is a nonlinear error correction mechanism (NEC) with linear cointegration given by $z_t^* = y_t - \gamma_1^* x_t$. For simplicity we impose the common factors restriction $\delta_1^* = 0$ on [4.4]

$$\Delta y_t = \gamma_1^* \Delta x_t + f(z_{t-1}^*, \gamma_2^*) + v_t \quad [4.5]$$

The errors of the cointegrating relation are given by $z_{t-1}^* = y_{t-1} - \gamma_1^* x_{t-1}$, and the OLS estimated residuals are given by $z_{t-1}^T = y_{t-1} - \gamma_1^T x_{t-1}$, where γ_1^T is the value of β estimated

in the regression $y_t = \alpha + \beta x_t + \varepsilon_t$, because $y_t = \gamma_1^* x_t + (z_t^0 + \mu)$, where $z_t = z_t^0 + \mu$ and z_t^0 is zero mean.

The analysis begin by generating the series z_t^* from [4.2]. If we differentiate in [4.2] with respect to z_{t-1}^* we obtain

$$\frac{d}{dz_{t-1}^*} z_t^* = 1 + \frac{d}{dz_{t-1}^*} f(z_{t-1}^*, \gamma_2^*) \quad [4.6]$$

and our boundedness condition is $-1 < \frac{d}{dz_{t-1}^*} z_t^* < 1$ which from Theorem 2.7 is sufficient to ensure that the series z_t^* is NED. In case that the non linear function is given by a more general nonlinear autoregressive model, the more general boundedness condition is explained in Mira and Escribano (1995).

Consider the exponential smooth transition error correction (ESTR-EC) model given by

$$f(z_t, \beta_1, \beta_2, \gamma_2) = \gamma_2 (1 - \exp(\beta_1(z_t - \beta_2)^2))$$

then the derivative [4.6] is $1 - 2\gamma_2 \beta_1(z_t - \beta_2) \exp(\beta_1(z_t - \beta_2)^2)$, and in order to verify the boundedness condition it has to occur that the sign of $\gamma_2 \beta_1(z_t - \beta_2) < 0$, which is a quite unattractive condition since in that case [4.2] generates appropriate series only for deviations from the equilibrium that are either positive or negative.

Consider now the logistic smooth transition error correction (LSTR-EC) model given by

$$f(z_t, \beta_1, \beta_2, \gamma_2) = \gamma_2 (1 + \exp(\beta_1(z_t - \beta_2)))^{-1}.$$

where the derivative [4.6] is $1 - \gamma_2 (\beta_1 \exp(\beta_1(z_t - \beta_2)))(1 + \exp(\beta_1(z_t - \beta_2)))^{-2}$. Figure 1 represents this derivative for the values of the parameters $(\beta_1, \beta_2, \gamma_2)$ given by $(8, 0, 1)$, $(6, 0, 1)$, $(3, 0, 1)$ and $(1, 0, 1)$ respectively. The shape of the derivative [4.6] depends on the values of the parameters considered. The boundedness condition requires that the graphs of the derivative are bounded by 1 in absolute value. As we see in the graphs those values approaches 1 asymptotically, then the condition holds but the behaviour of the series approaches the behaviour of a unit root process or of a "stable" process depending on the region where s takes values. We will come back to this point later on.

4.2 The Models

We propose three alternative models to generate the series $z_t^* \equiv s_t$, given by the following nonlinear functions.

Model 1: Take

$$f(s, \beta_1, \beta_2, \gamma_2) = -\gamma_2 \arctan(\beta_1 s + \beta_2)$$

for $\gamma_2 > 0$.

Model 2: Take

$$f(s, \beta_1, \beta_2, \beta_3, \beta_4, \gamma_2) = \gamma_2 (\exp(-\beta_1 s) - \beta_2) I_{\{s \geq 0\}} + \gamma_2 (\beta_4 - \exp(\beta_3 s)) I_{\{s < 0\}}$$

for $I_{\{Z\}}$ is the characteristic function of the set Z , $\gamma_2 > 0$, $\beta_1 > 0$ and $\beta_3 > 0$.

Model 3: Take

$$f(s, \beta_1, \beta_2, \beta_3, \beta_4, \gamma_2) = -\gamma_2 ((s + \beta_1)^3 + \beta_2) / ((s + \beta_3)^2 + \beta_4)$$

for $\gamma_2 > 0$.

In the three cases the derivatives are in the desired region for the appropriate values of the parameters but not for others. The derivative [4.6] for Models 1, 2 and 3 are respectively

$$\begin{aligned} 1 & - \gamma_2 \frac{\beta_1}{1 + (\beta_1 s + \beta_2)^2} \\ 1 & + \gamma_2 (-\beta_1 \exp(-\beta_1 s)) I_{\{s \geq 0\}} + \gamma_2 (-\beta_3 \exp(\beta_3 s)) I_{\{s < 0\}} \\ 1 & - \gamma_2 \frac{3(s + \beta_1)^2((s + \beta_3)^2 + \beta_4) - 2(s + \beta_3)((s + \beta_1)^3 + \beta_2)}{((s + \beta_3)^2 + \beta_4)^2} \end{aligned}$$

In particular in Model 1 for values of $(\beta_1, \beta_2, \gamma_2)$ equal to $(1, 0, 1)$ the derivative is $1 - \frac{1}{1+s^2}$ which is always between 0 and 1, but this is not necessarily true for other parameter values. In Model 2, for values of $(\beta_1, \beta_2, \beta_3, \beta_4, \gamma_2)$ equal to $(1, 0, 1, 0, 1)$ the derivative is $1 - \exp(-s) I_{\{s \geq 0\}} - \exp(s) I_{\{s < 0\}}$. In Model 3 for values of $(\beta_1, \beta_2, \beta_3, \beta_4, \gamma_2)$ equal to $(0, 0, 0, 1, 1)$ the function $f(s)$ is $-\frac{s^3}{s^2+1}$ which has an asymptote in the line $y = -x$ and is above it in the positive part and below it in the negative part. Therefore the derivative will be always smaller than 1 and will approach 0 asymptotically. Figures 2 to 3 present some examples for other parameter values. Figure 2 present the derivative $1 + f'(s)$ for Model 2 where the values $(\beta_1, \beta_3, \gamma_2)$ are given by $(1, 1, 1)$, $(1, 5, 0.5)$, $(4, 4, 1)$, $(1, 1, 1.5)$. The comments we made for Figure 1 apply now. Figure 3 presents the derivative for Model 3 where the values $(\beta_1, \beta_2, \beta_3, \beta_4, \gamma_2)$ are given by $(0, 0, 0, 1, 1)$, $(0.5, 0, 0, 1, 0.7)$, $(0, 0, 1, 1, 0.7)$, $(0, 0, 0, 5, 0.7)$. Within this class of rational polynomials the models considered satisfy the condition on the absolute value of the derivative, except maybe for a small region of values of s . Therefore, we feel more confident estimating these type of NEC models.

The following table provides the values of the parameters for Models 1, 2 and 3 that will be analyzed in the Monte Carlo simulations.

Table 0	Parameters	Values		
Model 1	$(\beta_1, \beta_2, \gamma_2)$	(2.0, 0.7)		
Model 2	$(\beta_1, \beta_2, \beta_3, \beta_4, \gamma_2)$	(1.1, 5, 5, 0.5)	(1.1, 1.1, 1.1)	
Model 3	$(\beta_1, \beta_2, \beta_3, \beta_4, \gamma_2)$	(0, 0, 0, 1, 1)	(0.5, 0, 0, 1, 0.7)	(0, 0, 1, 1, 0.7)

We should remember that case 2 of Model 2 and case 1 of Model 3 do not verify the boundedness condition [4.6]. Since the bound is exceeded in a small enough region that might allows the series to be NED in practice, and we will study this possibility later on.

The effect on z_t^* produced by the nonlinear part in the series is significantly different from the effects produced by linear models. The following figures provides some examples. Figures 4 to 7 present (in rows) the plots of the time series z_t^* and $f(z_{t-1}^*, \gamma_2)$ as well as the diagram of z_t^* vs. $f(z_{t-1}^*, \gamma_2)$, and the histogram of z_t^* . Figure 4 presents Model LSTR-EC with values of $(\beta_1, \beta_2, \gamma_2)$ given by (6, 0, 1). Figure 5 present Model 2, and Figures 6 and 7 present Model 3. In all those cases the values of the parameters are given in the above Table.

There are four graphs, (1,1), (1,2), (2,1) and (2,2), inside each figure which provide information on some aspects of the series generated with the models. Graphs (1,1) and (2,2) of Figures 5 to 7 show asymmetric behaviour in the series. Graphs (1,2) shows that the action of the nonlinear part can be frequent or occasional, or asymmetric in different senses. This behaviour produces values that appear as outliers, as in Figure 7. Graph (2,1) shows that the nonlinear adjustment can be symmetric, asymmetric or very asymmetric.

The parametric error correction functions proposed in this paper are illustrated in Graphs (2,1) of Figures 4 to 7. Those graphs show the adjustment of the nonlinear error correction mechanisms $f(\cdot)$ given in [4.5]. In all cases the adjustment $f(z_{t-1})$ assigns a positive correction to a negative error z_{t-1} and conversely. In almost every case the adjustment is error correcting ("stable") in the sense that (besides the latter condition) it holds that $|f(z_{t-1})| < |z_{t-1}|$. In many cases the correction is proportional to the error, i.e. $|f(z)/z|$ growth with $|z|$, see Figures 6 and 7. In this sense Model 3 is better than Models 1, 2 and Model LSTR-EC.

The third set of figures present the series y_t and x_t as well as its cointegration error z_t^* , for the value $\gamma_1^* = 0.7$. The Figures 8 to 10 present (in rows) four series: the equilibrium errors z_t^* given in [4.2], the series x_t given in [4.1], the series y_t given in [4.3], and the partial sum of the equilibrium errors z_t^* . Figure 8 presents Model 2.2 and Figures 9 and 10 present Model 3.1 and 3.3, where the values of the parameters are given above. A conclusion comes out clearly, in all cases z_t^* looks like I(0), see Graphs (1,1), and its partial sum like I(1), with and without drift, see Graphs (2,2).

4.3 Biases in the Two-Step Estimator

The following tables show the bias obtained in the OLS estimation of the value of the cointegrating parameter γ_1^* in [4.3] from the regression in levels, and the biases obtained by the OLS estimation of the rest of the parameters of [4.5] (step 2). Within step 2 we distinguish two cases: step 2a analyze the bias of the short run parameters (γ_1 and γ_2) obtained by assuming that the cointegrating vector γ_1^* is known and step 2b assumes that it is obtained by OLS $\hat{\gamma}_1$ from step 1. The estimation is OLS because in [4.5] both parameters γ_1 and γ_2 enter linearly. Notice the fact that we estimate γ_1 in the second step is an implication of the imposed common factor restriction. No lack of generality is implied by this restriction since if we do not impose it a different value, say $\gamma_1^* + \lambda^*$, would be estimated. The subcases correspond to sample sizes $T=100$ and $T=200$ respectively (with the first 50 observations discarded from samples of 150 and 250). The estimation is done with $N=1000$ replications and the results given in the tables are the mean given by $\bar{\gamma}^T = \frac{1}{N} \sum_{i=1}^N (\gamma_i^T - \gamma^*)$ and the standard deviation given by $\sqrt{\frac{1}{N} \sum_{i=1}^N (\gamma_i^T - \bar{\gamma}^T)^2}$ Table 1 present the results for Model 1.

Table 1	Step 1	Step 2a		Step 2b	
	γ_1	γ_1	γ_2	γ_1	γ_2
T=100	-0.003757	-0.000697	0.017341	-0.002220	0.048512
(a)	(0.059096)	(0.104932)	(0.104655)	(0.109761)	(0.308465)
T=200	-0.000250	0.002261	0.008467	0.002114	0.011559
(a)	(0.030922)	(0.071985)	(0.071539)	(0.074518)	(0.077149)
(a) The values of the parameters are: $\gamma_1^* = 0.7$ and $\gamma_2^* = 0.7$					

Table 2 present the results for Model 2.

Table 2	Step 1	Step 2a		Step 2b	
	γ_1	γ_1	γ_2	γ_1	γ_2
T=100	-0.001414	-0.000067	-0.030597	-0.000966	-0.911666
(a)	(0.068766)	(0.102604)	(0.169106)	(0.106581)	(0.34388)
T=200	-0.000234	0.002318	-0.019718	0.002225	-0.034083
(a)	(0.034872)	(0.072017)	(0.119417)	(0.074083)	(0.130607)
T=100	0.006384	0.000194	-0.004509	0.000778	-0.309411
(b)	(0.10633)	(0.103548)	(0.05413)	(0.119821)	(2.674576)
T=200	0.00026	0.002668	-0.001408	0.002195	-0.117544
(b)	(0.056779)	(0.073446)	(0.03892)	(0.080578)	(2.650578)
(a) The values of the parameters are: $\gamma_1^* = 0.7$ and $\gamma_2^* = 0.5$					
(b) The values of the parameters are: $\gamma_1^* = 0.7$ and $\gamma_2^* = 1$					

Table 3 present the results for Model 3.

Table 3	Step 1	Step 2a		Step 2b	
	γ_1	γ_1	γ_2	γ_1	γ_2
T=100	-0.0007091	0.000063	0.019754	0.000350	0.002616
(a)	(0.048779)	(0.102697)	(0.126568)	(0.107228)	(0.139435)
T=200	0.000643	-0.000756	0.015377	-0.0005855	0.008316
(a)	(0.024303)	(0.075611)	(0.093610)	(0.077507)	(0.098883)
T=100	-0.0005	0.001941	0.029052	0.00197	0.01316
(b)	(0.063709)	(0.101769)	(0.114767)	(0.106327)	(0.140009)
T=200	0.000724	-0.00042	0.019576	-0.000273	0.0108464
(b)	(0.031319)	(0.073394)	(0.080676)	(0.074955)	(0.095423)
T=100	0.005497	0.005394	0.021408	0.006727	-0.007028
(c)	(0.064896)	(0.098307)	(0.097784)	(0.103658)	(0.164639)
T=200	0.000787	-0.000864	0.008801	-0.000625	0.00229
(c)	(0.032529)	(0.07554)	(0.069158)	(0.077683)	(0.108525)
(a) The values of the parameters are: $\gamma_1^* = 0.7$ and $\gamma_2^* = 1$					
(b) The values of the parameters are: $\gamma_1^* = 0.7$ and $\gamma_2^* = 0.7$					
(c) The values of the parameters are: $\gamma_1^* = 0.7$ and $\gamma_2^* = 0.7$					

In the second step the biases are sistematically greater with $\hat{\gamma}_1$ (2b) than with γ_1^* (2a), except in some cases of Model 3 (but in those cases the difference is not significant, given the obtained standard deviation). The biases are about 1% of the size of the parameters in both cases. The bias is always greater for the estimation of γ_2^* than for the estimation of γ_1^* .

Now we present the simulation results of model [4.5] when the estimation procedure is nonlinear least squares (NLS). The procedure used is the *S-plus* function *ms*(\cdot). The sample sizes are T=100, T=500 and T=1000 and the initial 100 observations have been disregared. The parameter γ_1^* takes always the value 0.7 and the initial value in the iterations of the calculation of *ms* is 1.

Table 4 shows the mean and the standard deviation for Model 1 with $(\beta_1, \beta_2, \gamma_2)$ equal to (2.0, 1). The initial values are (1, 1, 1).

Table 4	γ_1	γ_2	β_1	β_2
T=100	-0.004934	0.1782	-133.386	6.77
(a)	(0.10478)	(0.70808)	(2890.57)	(318.079)
T=500	0.00199	0.0150	-0.10739	-0.00534
(a)	(0.04448)	(0.1144)	(0.585)	(0.159)
T=1000	-0.00184	0.00631	-0.04648	0.00221
(a)	(0.0307)	0.0786	0.38279	(0.10068)
(a) The values of the parameters are given in Table 0, Model 1				

Table 5 shows the results for Model 2 with $(\beta_1, \beta_2, \beta_3, \beta_4, \gamma_2)$ equal to $(1, 1, 1, 1, 1)$ and $(2, 2, 1, 1, 1)$. The initial values are $(0.5, 0.5, 0.5, 0.5, 1)$.

Table 5	γ_1	γ_2	β_1	β_2	β_3	β_4
T=100	-0.001303	-0.76778	-1.78355	-0.1409	0.60924	-0.2044
(a)	(0.10177)	(1.5968)	(8.1002)	(1.6105)	(0.3633)	(1.4762)
T=500	0.000713	-0.27182	-1.3345	0.0999	0.7273	-0.2272
(a)	(0.04477)	(0.73446)	(4.59)	(0.7266)	(0.1258)	(0.62104)
T=1000	0.001037	-0.14254	-1.3786	0.12322	0.7342	-0.2338
(a)	(0.03130)	(0.5427)	(3.511)	(0.6687)	(0.08262)	(0.5750)
T=100	0.00432	-0.5679	-0.166	0.2004	0.5019	-0.3895
(b)	(0.0994)	(1.094)	(4.882)	(1.2065)	(0.4892)	(0.605)
T=500	-0.00078	-0.04238	-0.08049	0.10269	0.6011	-0.48535
(b)	(0.04409)	(0.45038)	(2.7578)	(0.8736)	(0.2151)	(0.5189)
T=1000	0.001094	0.0835	-0.4141	0.1387	0.6021	-0.5622
(b)	(0.03243)	(0.2595)	(2.465)	(0.731)	(0.1386)	(0.475)
(a) The values of the parameters are given in Table 0, Model 2.1						
(b) The values of the parameters are given in Table 0, Model 2.2						

Finally, Table 6 presents the results for Model 3. The values of $(\beta_1, \beta_3, \beta_4, \gamma_2)$ are $(0, 0, 1, 1)$, $(0.5, 0, 1, 0.7)$, $(0, 1, 1, 0.7)$, and the initial values are $(0.5, 0.5, 0.5, 1)$.

Table 6	γ_1	γ_2	β_1	β_3	β_4
T=100	0.00035	0.87329	-0.12535	-0.24871	-3.5068
(a)	(0.10395)	(15.542)	(0.5599)	(1.39106)	(66.9984)
T=500	-0.0004	0.0042	-0.004656	-0.002859	-0.03609
(a)	(0.04760)	(0.14929)	(0.25903)	(0.43392)	(0.5735)
T=1000	0.00285	0.0052	-0.001516	-0.00121	-0.02103
(a)	(0.03187)	(0.09239)	(0.16617)	(0.28654)	(0.3542)
T=100	0.001803	0.86532	0.55183	-0.52427	-3.68548
(b)	(0.10652)	(11.5195)	(0.67815)	(2.08455)	(45.14084)
T=500	0.000059	0.010005	0.92073	-0.03725	-0.09512
(b)	(0.046)	(0.15727)	(0.36917)	(0.48242)	(0.920102)
T=1000	-0.00019	0.0071	0.9772	0.0014	-0.07278
(b)	(0.0317)	(0.1047)	(0.2470)	(0.34016)	(0.6141)
T=100	-0.00027	3.525	-0.135	2.073	-21.084
(c)	(0.1036)	(25.243)	(0.6393)	(3.4123)	(182.786)
T=500	-0.00098	0.04306	-0.00723	1.972	-0.2902
(c)	(0.0447)	(0.1847)	(0.2414)	(0.2477)	(1.197)
T=1000	0.00000	0.01604	-0.00447	1.986	-0.1054
(c)	(0.0326)	(0.1001)	(0.1618)	(0.1655)	(0.6109)
(a) The values of the parameters are given in Table 0, Model 3.1					
(b) The values of the parameters are given in Table 0, Model 3.2					
(c) The values of the parameters are given in Table 0, Model 3.3					

As can be seen the minimum sample size to estimate with small biases has to be closer to 500 than to 100 observations. With T=500 the bias is about 1% of the size of the parameter and also the standard deviations. The greatest biases always correspond to the nonlinear parameters. In Model 1 there are expectacular decreases in the biases as we go from 100 to 500 observations. and the same occurs with β_4 in Model 3. For Model 2 the biases are smaller but more persistent and still are significant for T=500.

5 Conclusions

There are evidence of interesting empirical macroeconomic applications where nonlinearities are found in an error correction context. However, no statistical results are available that justify parameter estimation and inference within this nonlinear context. To start filling this gap. we extend certain results of linear integrated and cointegrated variables to a nonlinear

framework, by introducing a concept of integration based on near epoch dependence requirements. Within this framework we are able to generalize certain properties of Granger's representation theorem to the nonlinear case. We found that if the variables are $I(1)$ with a nonlinear error correction system then they are linearly cointegrated, under certain conditions on the nonlinear functions. In particular, we give sufficient conditions for the NEC to be well specified and balanced.

Furthermore we study the consistency of the least squares estimator of linear or nonlinear cointegrating relationship. We derive the asymptotic distribution of the least squares estimator of the rest of the parameters of the dynamic NEC model. In particular, our conditions allow us to extend the two-step least squares (LS) approach of Engle and Granger(1987) to a nonlinear context (NLS). We graphically analyze the properties of NEC models showing that our sufficient condition on the nonlinear function is satisfied for certain parameter values of the functional forms, like the logistic, switching exponential, arctangent and rational polynomials, but not necessarily for others parameter values. We suggest to graphically represent the first order derivative and when this condition is violated in certain range but not always we suggest to check for the series to be $I(0)$ and for its partial sum to be $I(1)$. Furthermore, we show how simple nonlinear error correction models can generate asymmetric time series behavior. We conclude that among the previously mentioned functional forms considered, the rational polynomials are the only ones that have the attractive property of correcting fast large equilibrium errors.

The small sample properties of the two-step least squares estimator are investigated by Monte Carlo simulations. We found that when the cointegrating relationship is linear but the error correction is nonlinear, the biases of the OLS estimator of the cointegrating vector are small in samples of sizes 100 and 200. The same results are maintained when the second step of the NEC model can be estimated by OLS, as in the empirical example of the money demand of the U.K from 1878-1970. However, the results are not as good when the NEC in the second step has to be estimated by nonlinear least squares (NLS). In that case the biases of the parameters that enter linearly in the models are small for sample sizes of 100, but we need sample sizes of 500 specially for those parameters that enter linearly but are related to the NEC term. With respect to the parameters of the nonlinear function that enter nonlinearly, the results of NLS depends on the particular parameter values of the DGP considered and on the functional form chosen, but in general we have to go to sample sizes near 1000 to have confidence on having small biases in all of the parameters.

A Appendix to Section 2

A.1 The $\|\cdot\|_S$ Norm

The matrix norm $\|\cdot\|_S$ is defined as follows

$$\|A\|_S \equiv \|(MD_\delta)^{-1}A(MD_\delta)\|_\infty$$

for M and D_δ being matrices that depend on the matrix A . Analogously the asociated vectorial norm is

$$\|Y\|_S \equiv \|(MD_\delta)Y\|_\infty$$

The norm $\|\cdot\|_S$ is characterized by the fact that

$$\|A\|_S \leq \rho(A) + \delta$$

for $\rho(A)$ being the spectral radious of A . See Mira and Escibano (1995).

A.2 Proof of Theorem 2.6

Define

$$\bar{Z}_t = \begin{cases} F(\bar{Z}_{t-1}) & \text{for } t > 0 \\ 0 & \text{for } t \leq 0 \end{cases}$$

and

$$\tilde{Z}_{t,s}^m = \begin{cases} \Phi \tilde{W}_{t-1} + F(\tilde{Z}_{t-1,s+1}^m) + U_t & \text{for } s+1 \leq m \\ \bar{Z}_t & \text{for } s+1 > m \end{cases}$$

where $\tilde{W}_t = E(W_t | A_t, \dots, A_{t-m})$, and therefore $E\|W_t - \tilde{W}_t\|_S^2 \leq \psi_m$ such that $\psi_m \rightarrow 0$ when $m \rightarrow \infty$. Then it is clear that $\tilde{Z}_{t,0}^m$ is $\sigma(U_t, \tilde{W}_{t-1}, \dots, U_{t-m+1}, \tilde{W}_{t-m})$ -medible, and then it is $\sigma(U_t, A_{t-1}, \dots, U_{t-m+1}, A_{t-m}, \dots, A_{t-2m})$ -medible,

The difference between Z_t and its predictor \bar{Z}_t is bounded for $t > 0$, because

$$\begin{aligned} \|Z_t - \bar{Z}_t\|_S &= \|\Phi W_{t-1} + F(Z_{t-1}) + U_t - F(\bar{Z}_{t-1})\|_S \\ &\leq \|\Phi W_{t-1} + U_t\|_S + \|F(Z_{t-1}) - F(\bar{Z}_{t-1})\|_S \end{aligned}$$

and by the Mean Value Theorem

$$\begin{aligned}
F(Z_t) - F(\bar{Z}_t) &= \begin{pmatrix} F_1(Z_t) - F_1(\bar{Z}_t) \\ \vdots \\ F_r(Z_t) - F_r(\bar{Z}_t) \end{pmatrix} \\
&= \begin{pmatrix} \frac{\partial F_1}{\partial z_1}(\ddot{Z}_t)(z_{1t} - \bar{z}_{1t}) + \cdots + \frac{\partial F_1}{\partial z_r}(\ddot{Z}_t)(z_{rt} - \bar{z}_{rt}) \\ \vdots \\ \frac{\partial F_r}{\partial z_1}(\ddot{Z}_t)(z_{1t} - \bar{z}_{1t}) + \cdots + \frac{\partial F_r}{\partial z_r}(\ddot{Z}_t)(z_{rt} - \bar{z}_{rt}) \end{pmatrix} \\
&= \begin{pmatrix} \frac{\partial F_1}{\partial z_1}(\ddot{Z}_t) & \cdots & \frac{\partial F_1}{\partial z_r}(\ddot{Z}_t) \\ \vdots & & \vdots \\ \frac{\partial F_r}{\partial z_1}(\ddot{Z}_t) & \cdots & \frac{\partial F_r}{\partial z_r}(\ddot{Z}_t) \end{pmatrix} \begin{pmatrix} z_{1t} - \bar{z}_{1t} \\ \vdots \\ z_{rt} - \bar{z}_{rt} \end{pmatrix} \\
&= \nabla_Z F(\ddot{Z}_t)(Z_t - \bar{Z}_t).
\end{aligned}$$

Now, since $\|\cdot\|_S$ is a subordinate matrix norm we have that

$$\begin{aligned}
\|Z_t - \bar{Z}_t\|_S &\leq \|\Phi\|_S \|W_{t-1}\|_S + \|U_t\|_S + \|\nabla_Z F(\ddot{Z}_t)\|_S \|(Z_{t-1} - \bar{Z}_{t-1})\|_S \\
&\leq \delta_{WU,t} + \delta_Z \|(Z_{t-1} - \bar{Z}_{t-1})\|_S
\end{aligned}$$

for some $N_{WU,t}$ and since $Z_0 = \bar{Z}_0 = 0$, then by iteration we obtain

$$\begin{aligned}
\|Z_t - \bar{Z}_t\|_S &\leq \sum_{j=0}^{t-1} N_{WU,t-j} \delta_Z^j \\
\|Z_t - \bar{Z}_t\|_S^2 &\leq \sum_{j=0}^{t-1} N_{WU,t-j}^2 \delta_Z^{2j} + \sum_{j=0}^{t-1} \sum_{i \neq j}^{t-1} N_{WU,t-i} N_{WU,t-j} \delta_Z^{i+j} \\
E\|Z_t - \bar{Z}_t\|_S^2 &\leq \Delta_{Z-\bar{Z}}^{(2)}
\end{aligned}$$

for some bound $\Delta_{Z-\bar{Z}}^{(2)}$, because, for instance, $E(N_{WU,t}) = \|\Phi\|_S \Delta_W^{(1)} + \Delta_U^{(1)}$. Now,

$$\begin{aligned}
\|Z_t - \tilde{Z}_{t,0}^m\|_S &= \|\Phi W_{t-1} + F(Z_{t-1}) + U_t - \Phi \tilde{W}_{t-1} - F(\tilde{Z}_{t-1,1}^m) - U_t\|_S \\
&\leq \|\Phi(W_{t-1} - \tilde{W}_{t-1})\|_S + \|F(Z_{t-1}) - F(\tilde{Z}_{t-1,1}^m)\|_S
\end{aligned}$$

and again by the Mean Value Theorem we obtain

$$\|F(Z_{t-1}) - F(\tilde{Z}_{t-1,1}^m)\|_S \leq \|\nabla_Z F(Z)\|_S \|Z_{t-1} - \tilde{Z}_{t-1,1}^m\|_S,$$

and since $\|\nabla_Z F(Z)\|_S \leq \delta_Z$ we have

$$\|Z_t - \tilde{Z}_{t,0}^m\|_S \leq \|\Phi\|_S \|W_{t-1} - \tilde{W}_{t-1}\|_S + \delta_Z \|Z_{t-1} - \tilde{Z}_{t-1,1}^m\|_S$$

and by iteration

$$\|Z_t - \tilde{Z}_{t,0}^m\|_S \leq \sum_{i=0}^m \delta_Z^i \|\Phi\|_S \|W_{t-1-i} - \tilde{W}_{t-1-i}\|_S + \delta_Z^m \|Z_{t-m} - \bar{Z}_{t-m}\|_S$$

and taking expectations

$$\begin{aligned} E\|Z_t - \tilde{Z}_{t,0}^m\|_S^2 &\leq E\left(\sum_{i=0}^m \delta_Z^i \|\Phi\|_S \|W_{t-1-i} - \tilde{W}_{t-1-i}\|_S\right)^2 + \delta_Z^{2m} E\|Z_{t-m} - \bar{Z}_{t-m}\|_S^2 \\ &\quad + 2E\left(\left(\sum_{i=0}^m \delta_Z^i \|\Phi\|_S \|W_{t-1-i} - \tilde{W}_{t-1-i}\|_S\right) \times \delta_Z^m \|Z_{t-m} - \bar{Z}_{t-m}\|_S\right). \end{aligned}$$

If we use for the third term in the summation the Holder inequality with $p = \frac{1}{2} = q$, i.e., $E|YX| \leq E^{1/2}|Y|^2 + E^{1/2}|X|^2$, only remains to work out the following term

$$\begin{aligned} &E\left(\sum_{i=0}^m \|W_{t-1-i} - \tilde{W}_{t-1-i}\|_S\right)^2 \\ &= E\sum_{i=0}^m \|W_{t-1-i} - \tilde{W}_{t-1-i}\|_S^2 \\ &\quad + E\sum_{i=0}^m \sum_{j \neq i}^m \|W_{t-1-i} - \tilde{W}_{t-1-i}\|_S \|W_{t-1-j} - \tilde{W}_{t-1-j}\|_S \\ &\leq \sum_{i=0}^m E\|W_{t-1-i} - \tilde{W}_{t-1-i}\|_S^2 \\ &\quad + \sum_{i=0}^m \sum_{j \neq i}^m E^{1/2}\|W_{t-1-i} - \tilde{W}_{t-1-i}\|_S^2 E^{1/2}\|W_{t-1-j} - \tilde{W}_{t-1-j}\|_S^2 \end{aligned}$$

and since $E\|W_{t-1-i} - \tilde{W}_{t-1-i}\|_S^2 = \psi_m$ then $E\|Z_t - \tilde{Z}_{t,0}^m\|_S^2$ is bounded by a summation of terms with ψ_m or terms with δ_Z and since ψ_m goes to zero and $0 < \delta_Z < 1$ we obtain

$$\lim_{m \rightarrow \infty} E\|Z_t - \tilde{Z}_{t,0}^m\|_S = 0$$

Now, given $E_{t-2m}(Z_t) \equiv E(Z_t | U_t, A_{t-1}, \dots, U_{t-2m+1}, A_{t-2m})$, we can obtain a bound for $\|Z_t - E_{t-2m}(Z_t)\|_{L^2}$. Since $\tilde{Z}_{t,0}^m$ is $\sigma(U_t, \dots, U_{t-m+1}, A_{t-m}, \dots, A_{t-2m})$ -medible then it is $\sigma(U_t, \dots, U_{t-2m+1}, A_{t-2m})$ -medible so that

$$\begin{aligned} \|Z_t - E_{t-2m}(Z_t)\|_{L^2} &\leq \delta_K \|Z_t - \tilde{Z}_{t,0}^m\|_{L^2} \\ &= \delta_K E^{1/2} \|Z_t - \tilde{Z}_{t,0}^m\|_S^2 \end{aligned}$$

and since $E\|Z_t - \tilde{Z}_{t,0}^m\|_S^2 \rightarrow 0$ at exponential rate then $\{Z_t\}$ is NED on the underlying sequence

$\{(U_t, W_t)\}$ of any size. Note that the first inequality is a generalization of the well known fact $E|Z_t - E(Z_t|I_t)|^2 \leq E|Z_t - g(I_t)|^2$ for any function $g(\cdot)$ of the information set I_t and δ_Z is some constant that depends on the norm $\|\cdot\|_S$.

Q.E.D.

A.3 Proof of Theorem 2.7

For parts (i) and (ii), let us write

$$\begin{aligned} X_t &= F(X_{t-1}, X_{t-2}) + \varepsilon_t \\ &= G(X_{t-1}) + \Phi_2 X_{t-2} + \varepsilon_t \\ \Delta X_t &= G(X_{t-1}) - X_{t-1} + \Phi_2 X_{t-2} + \varepsilon_t \\ &= (-\Phi_2)(X_{t-1} - X_{t-2}) - (I - \Phi_2)X_{t-1} + G(X_{t-1}) + \varepsilon_t \\ &= \Psi_1 \Delta X_{t-1} + H(X_{t-1}) + \varepsilon_t \end{aligned}$$

where $\Psi_1 = -\Phi_2$, and $H(X_{t-1}) = -(I - \Phi_2)X_{t-1} + G(X_{t-1})$. Now, since ε_t and ΔX_t are $I(0)$ then $H(X_{t-1})$ is also $I(0)$, even though X_t is not. Part (iii) is immediate from Theorem 2.6. *Q.E.D.*

A.4 Sketch of the Proof of Theorem 2.8

Let us normalize the $(r \times n)$ matrix, base of the space of cointegration relations, in the following way $\alpha' = [I, -\beta']$ such that $\alpha'X_t = Z_t$, and let us define the $(n \times n)$ matrix M as

$$M = \begin{pmatrix} I & -\beta' \\ 0 & I \end{pmatrix}.$$

Then $MX_t = [Z_t', X_{2t}']'$ for some partition of the vector X_t as $X_t' = [X_{1t}', X_{2t}']$, with X_{1t} of dimension $(r \times 1)$ and X_{2t} of dimension $((n - r) \times 1)$. Given the NEC representation

$$\Delta X_t = \Psi \Delta X_{t-1} + J(\alpha' X_{t-1}) + \varepsilon_t$$

if we multiply by M we obtain the following system

$$\begin{aligned}\alpha' \Delta X_t &= \alpha' \Psi \Delta X_{t-1} + \alpha' J(\alpha' X_{t-1}) + \alpha' \varepsilon_t \\ \Delta X_{2t} &= \Psi_2 \Delta X_{t-1} + J_2(\alpha' X_{t-1}) + \varepsilon_{2t}\end{aligned}$$

for some partition of ε_t , Ψ , and $J(\alpha' X_{t-1})$. Let us represent the vector $[Z'_{t-1}, X'_{2,t-1}]'$ as L_{t-1} , then the system can be rewritten as

$$\begin{aligned}Z_t &= Z_{t-1} + \alpha' \Psi M^{-1} \Delta L_{t-1} + \alpha' J(\alpha' X_{t-1}) + \alpha' \varepsilon_t \\ \Delta X_{2t} &= \Psi_2 M^{-1} \Delta L_{t-1} + J_2(\alpha' X_{t-1}) + \varepsilon_{2t}\end{aligned}$$

or

$$\begin{aligned}Z_t &= Z_{t-1} + P \Delta L_{t-1} + K'(Z_{t-1}) + \eta_{1t} \\ \Delta X_{2t} &= \Psi_2 \Delta L_{t-1} + J_2(Z_{t-1}) + \eta_{2t}\end{aligned}$$

that is straightforward to rewrite as in [2.7].

Q.E.D.

B Appendix to Section 3

B.1 Proof of Proposition 3.1

Proof: Immediate from Theorems 3.3 and 3.7 of Escibano and Mira (1997) under the conditions of a FCLT for NED processes (see Theorem 2.1b).

B.2 Proof of Proposition 3.2

Proof: Once we know the cointegrating parameter α' we can apply Theorem 3.3 and Theorem 5.1 of Gallant and White (1988).

B.3 Proof of Theorem 3.4

Let us write [3.3b] as

$$G^T(\theta^*) = (F^*(\gamma_2^*) - F^T(\gamma_2^*)) + V. \quad [B.1]$$

The estimated value θ^T of θ^* for [B.1] verifies

$$\lim_{T \rightarrow \infty} T^{-1/2}(\theta^T - \theta^*)' = \lim_{T \rightarrow \infty} (T^{-1/2}G^T(\theta^*)'G_\theta^T(\theta^*)) (T^{-1}G_\theta^T(\theta^*)'G_\theta^T(\theta^*))^{-1}$$

and can be written as

$$\begin{aligned} \lim_{T \rightarrow \infty} T^{-1/2}(\theta^T - \theta^*)' &= \lim_{T \rightarrow \infty} (T^{-1/2}(F^*(\gamma_2^*) - F^T(\gamma_2^*))'G_\theta^T(\theta^*)) (T^{-1}G_\theta^T(\theta^*)'G_\theta^T(\theta^*))^{-1} \\ &\quad + \lim_{T \rightarrow \infty} (T^{-1/2}V'G_\theta^T(\theta^*)) (T^{-1}G_\theta^T(\theta^*)'G_\theta^T(\theta^*))^{-1} \end{aligned}$$

and we want that this limit equals the limit of the estimated value θ^T of θ^* for [3.3a]

$$\lim_{T \rightarrow \infty} T^{-1/2}(\theta^T - \theta^*)' = \lim_{T \rightarrow \infty} (T^{-1/2}G^*(\theta^*)'G_\theta^*(\theta^*)) (T^{-1}G_\theta^*(\theta^*)'G_\theta^*(\theta^*))^{-1}$$

then Assumption 3.3 ensures the equality.

Q.E.D.

References

- Balke, N. S. and T. B. Fomby, (1992), "Threshold Cointegration". *Working Paper, Economics Department, Southern Methodist University*.
- Bierens, H. J., (1981), *Robust Methods and Asymptotic Theory in Nonlinear Econometrics*. Lecture Notes in Economics and Mathematical Systems, vol. 192, Berlin: Springer Verlag.
- Burgess, S. M., (1992), "Nonlinear Dynamics in a Structural Model of Employment", *Journal of Applied Econometrics*, 7, S101-S118.
- Engle, R. F. and C. W. J. Granger, (1987), "Co-Integration and Error Correction: Representation, Estimation, and Testing", *Econometrica*, 55 (2), 251-276.
- Escribano, A., (1986). *Non-Linear Error-Correction: The Case of Money Demand in the U.K. (1878-1970)*. Chapter IV. Ph. D. Dissertation. University of California, San Diego.
- Escribano, A., (1987). "Error-Correction Systems: Nonlinear Adjustment to Linear Long-Run Relationships". *CORE Discussion Paper 8730. C.O.R.E.*
- Escribano, A. and S. Mira, (1997), "Nonlinear Cointegration", *Mimeo*.
- Gallant, A. R. and H. White, (1988), *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models*. New York: Basil Blackwell.
- Friedman, M. and A. Schwartz, (1982), *Monetary Trends in the United States and the United Kingdom: Their Relation to Income, Prices, and Interest Rates, 1867-1975*. Chicago: University of Chicago Press.
- Granger, C. W. J., (1995). "Modelling Nonlinear Relationships Between Extended-Memory Variables". *Econometrica*, 63 (2), 265-279.
- Granger, C. W. J., and J. Hallman, (1991), "Long Memory Series with Attractors", *Oxford Bulletin of Economics and Statistics*, 53, 11-26.
- Granger, C. W. J. and T. H. Lee, (1989), "Investigation of production, sales and inventory relationships using multicointegration and non-symmetric error correction models", *Journal of Applied Econometrics*, 4, S145-S159.
- Granger, C. W. J. and P. Newbold, (1974), "Spurious Regressions in Econometrics". *Journal of Econometrics*, 2, 111-120.

Granger, C. W. J., and N. Swanson, (1995). "Further developments in the study of cointegrated variables", *Mimeo*.

Granger, C. W. J., and T. Teräsvirta, (1993). *Modelling Nonlinear Economic Relationships*. New York: Oxford University Press.

Hendry, D. F. and N. R. Ericsson, (1991), "An Econometric Analysis of the U.K. Money Demand in "Monetary Trends in the United States and the United Kingdom" by Milton Friedman and Anna J. Schwartz", *The American Economic Review*, 81, 8-38.

Johansen, S., (1988). "Statistical Analysis of Cointegration Vectors", *Journal of Economics Dynamic and Control*, 12, 231-54.

Johansen, S., (1991), "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models", *Econometrica*, 59, 1551-1580.

Kunst, R. M., (1992). "Threshold Cointegration in Interest Rates", *Working Paper, Institute for Advanced Studies, Vienna*.

Longbottom, A. and S. Holly, (1985), "Econometric Methodology and Monetarism: Professor Friedman and Professor Hendry on the Demand for Money", *Discussion Paper No. 131. London Business School*.

Mira, S., (1996). *Modelos Económicos Dinámicos No Lineales con Tendencias Estocásticas*. Ph. D. Dissertation. Universidad Carlos III de Madrid. Madrid.

Mira, S. and A. Escribano, (1995). "Nonlinear Time Series Models: Consistency and Asymptotic Normality of NLS Under New Conditions". *Working Paper 95-42, Universidad Carlos III de Madrid. Madrid*.

Phillips, P. C. B., (1986). "Understanding Spurious Regressions in Econometrics", *Journal of Econometrics*. 33. 311-340.

Sims, C. A., (1972). "Money, Income, and Causality", *The American Economic Review*. 62. 540-552.

Stock, J. H., (1987). "Asymptotic Properties of Least Squares Estimation of Cointegrating Vectors". *Econometrica*. 55, 1035-1056.

Watson, M., (1994). *Large Sample Estimation and Hypothesis Testing*. Vector Autoregression and Cointegration. in "Handbook of Econometrics", vol. IV, eds. R.F. Engle and D. L. McFadden. Amsterdam: Elsevier Science

Wooldridge, J. M. and H. White, (1988), "Some Invariance Principles and Central Limit Theorems for Dependent Heterogeneous Processes", *Econometric Theory* 4, 210-30.

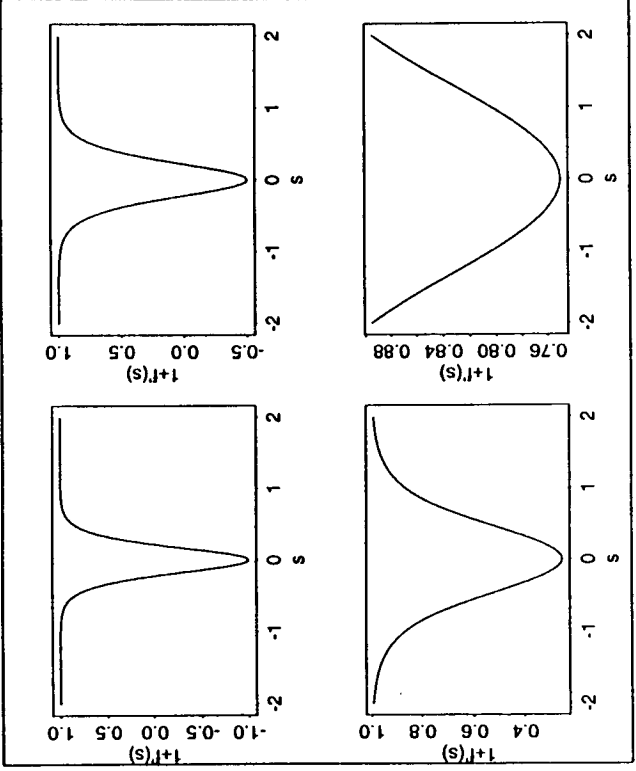


Figure 1: Model LSTR-EC

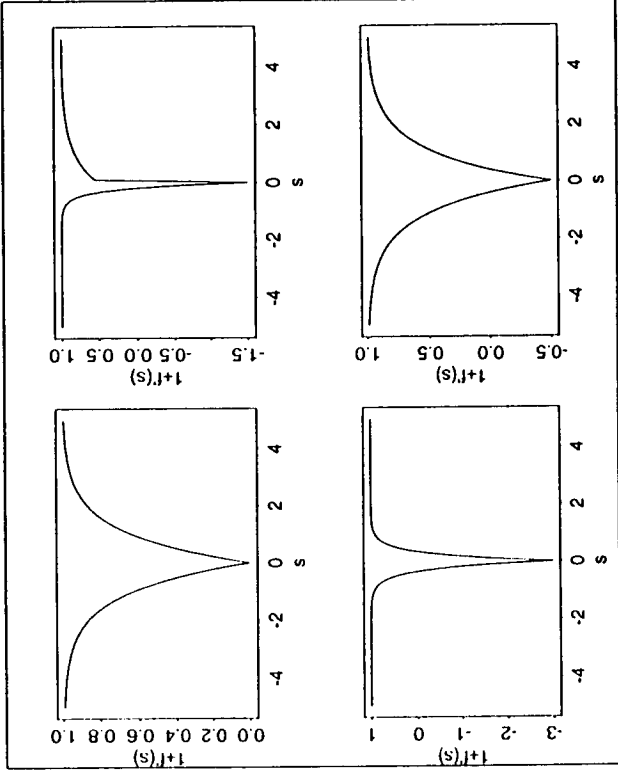


Figure 2: Model 2

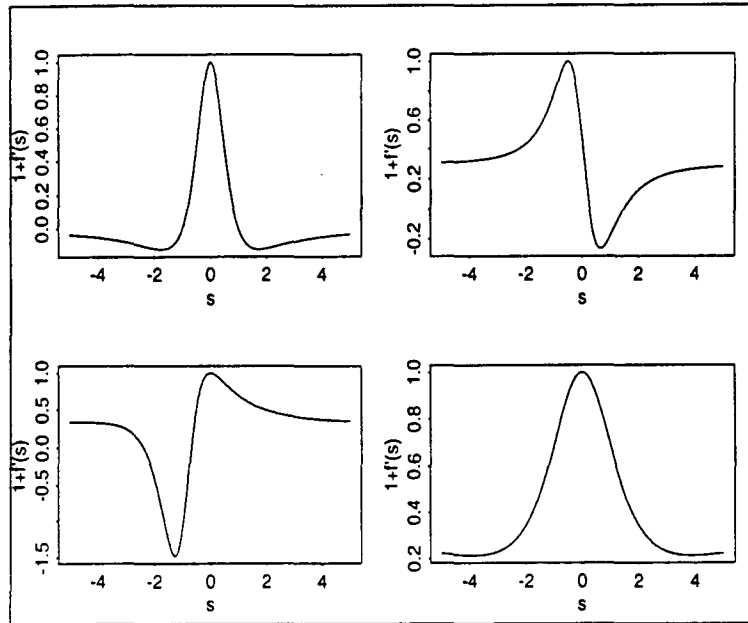


Figure 3: Model 3

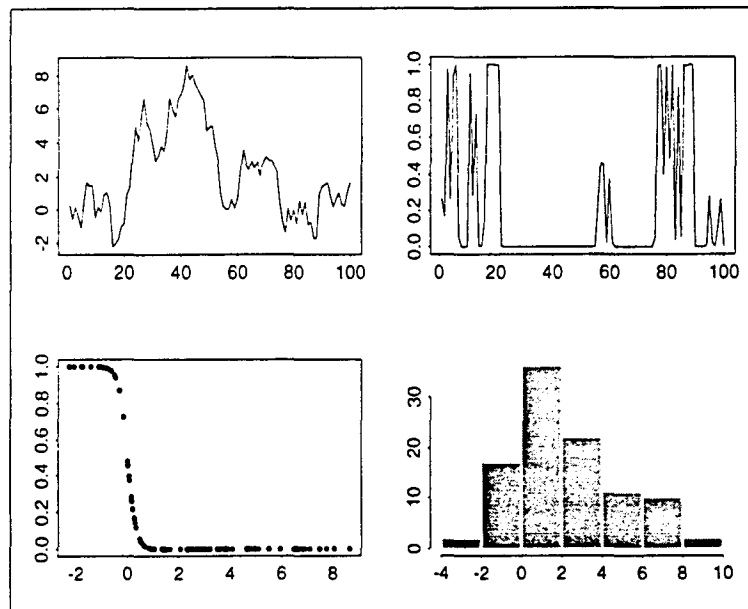


Figure 4: Model LSTR-EC

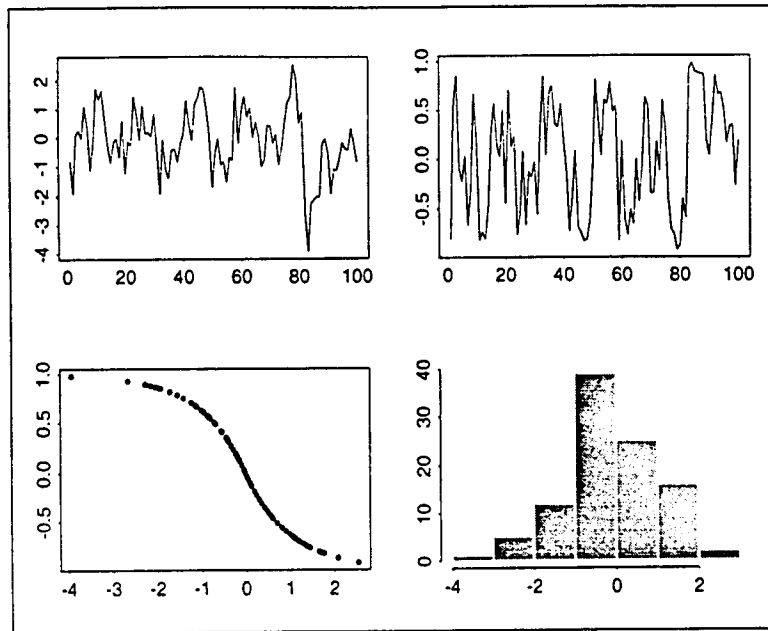


Figure 5: Model 2.1

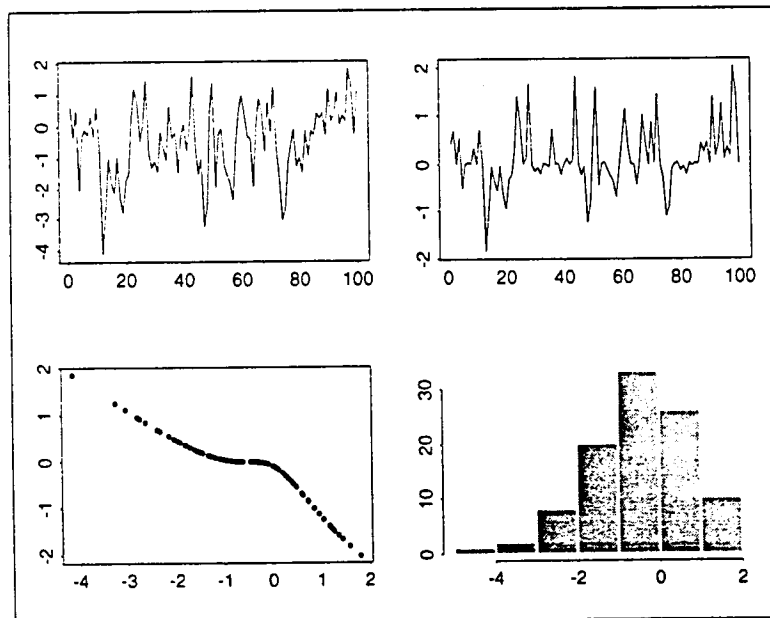


Figure 6: Model 3.2

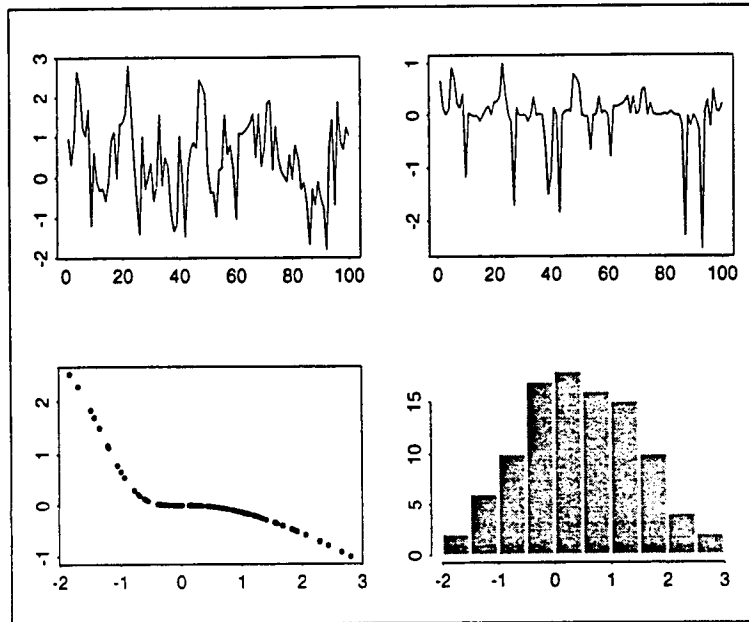


Figure 7: Model 3.3

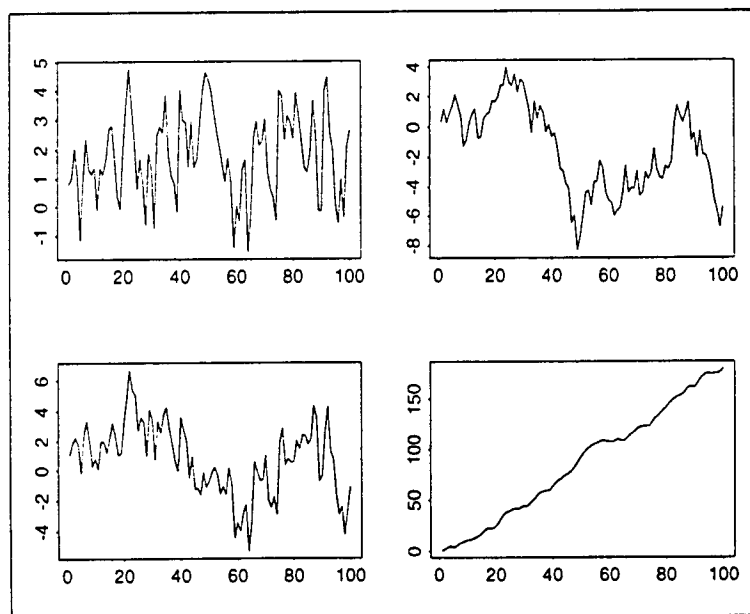


Figure 8: Model 2.2

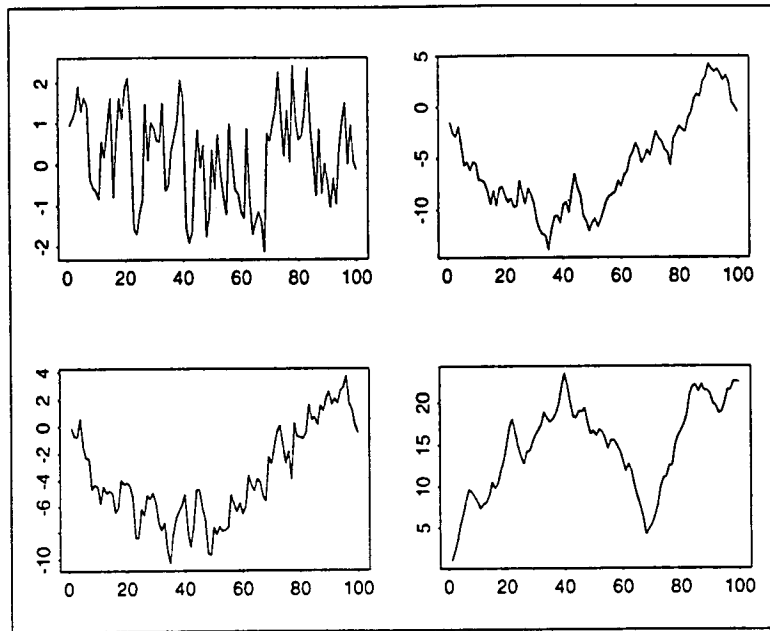


Figure 9: Model 3.1

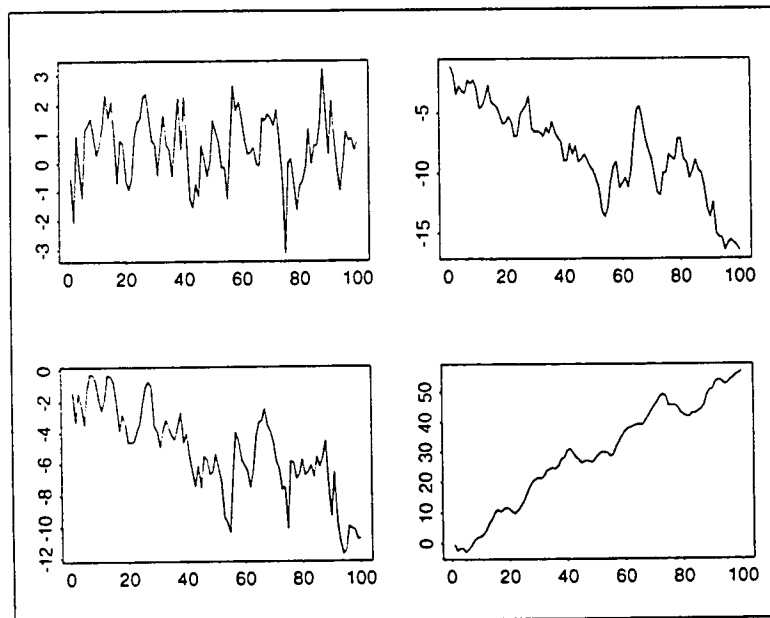


Figure 10: Model 3.3